

# Quality of Health IT Evaluations

Pirkko NYKÄNEN<sup>a,1</sup> and Johanna KAIPIO<sup>b</sup>

<sup>a</sup>*University of Tampere, School of Information Sciences, Tampere, Finland*

<sup>b</sup>*Aalto University, School of Science, Espoo, Finland*

**Abstract.** Health IT evaluation studies have often been found to be of limited quality. To address this problem, several guidelines and frameworks have been developed as tools to support improvement of the quality of evaluation studies. In this contribution, we review available guidelines and then present the Good Evaluation Practice Guideline in Health Informatics (GEP-HI) in more detail. GEP-HI is a comprehensive guideline which supports especially planning and execution of a health IT evaluation study. The GEP-HI guideline helps to overcome the quality problems related to weak study planning and methodological study design. We also discuss application of GEP-HI on an evaluation project and discuss the need to publish systematically following the recognised publication guidelines. Finally we discuss the future trend on multi-method evaluation approaches.

**Keywords.** Evaluation, quality, scope, guideline, framework, health IT.

## 1. Introduction

Evaluation is the means to assess the quality, value, effects and impacts of health IT in the health care environment. Evaluation is defined as the “act of measuring or exploring properties of a health information system, in planning, development, implementation, or operation, the result of which informs a decision to be made concerning that system in a specific context” [1, p. 480]. Evaluation offers methods and tools to collect evidence about the benefits, quality, effects and impacts of health IT.

In health care practice health IT applications offer challenging opportunities to improve the health care system’s functioning, effectiveness and outcomes as well as health care services quality and delivery, but there are also problems and unanticipated effects related to the use of IT [2,3].<sup>2</sup> It is of utmost importance in the health care environment that health IT provides the optimal and safe results and therefore health IT applications need to be evaluated with robust methodologies and evaluation results are to be reported following structured reporting standards [4,5].

Evaluation is difficult, it deals with values and norms and various organizational contexts and stakeholders’ interests, and it has to fight for funding and support [6]. Additionally, many potentially applicable methodologies and methods exist and evaluation results are to be analysed and interpreted in the study context [4]. A challenge to improve the quality of evaluation studies is to apply a systematic

<sup>1</sup> Corresponding author: Prof. Dr. Pirkko Nykänen, University of Tampere, School of Information Sciences, PO Box 607, FIN-33014 University of Tampere, Finland, Pirkko.Nykanen@uta.fi.

<sup>2</sup> See also: F. Magrabi et al., Health IT for patient safety and improving the safety of health IT, in: E. Ammenwerth, M. Rigby (eds.), *Evidence-Based Health Informatics*, Stud Health Technol Inform 222, IOS Press, Amsterdam, 2016.

approach, plan the evaluation study carefully and execute the study following systematic evaluation guidelines [7].

In this contribution we review the focus and scope of the published evaluation studies, discuss the quality problems related to these studies, and present health IT evaluation guidelines and frameworks, namely the GEP-HI guideline.

## **2. The scope and quality of evaluation studies**

The quality of an evaluation study is dependent on many factors, e.g. on the objectivity of the study and on the independence of evaluators, referring to their independence on economic interests, on intellectual interests and on the various stakeholders' interests. An evaluation study must also be scientifically well-established on robust theories and methodologies<sup>3</sup>, and the study should be performed following the principles of scientific research.

As there are many interesting focuses for evaluation, e.g. economics, efficiency, usability of health IT, safety, privacy and security, compliance with the clinical process, functionality of health IT, effects and impacts on health care outcomes, there is also need to use many different methods suitable for measuring the evaluation criteria of interest. These potential methods cover e.g. qualitative and quantitative methods, statistics, heuristics, ethnography, human-system interaction observations, data mining and quality analysis, cost-effectiveness and cost-benefit analyses.<sup>4</sup> There are also many potential perspectives for evaluation, representing various stakeholders' viewpoints, e.g. managerial, clinical, technical, and these viewpoints may be studied at various levels of health care system – local, regional and national, or even at EU or international levels.<sup>5</sup> It is of utmost importance, when planning an evaluation study, to elaborate and define how these issues are related to the study: Coverage of scientific robustness, relevance to the current purpose of the study, best fit of the important characteristics to the specific current need and relevance of the methods. The quality also comprises the quality in publication of the study.

Evaluation studies have been performed since 1960's, however, most with a rather narrow scope, often focusing on how health IT systems are related to professionals' roles, change management and user involvement [8]. Later the studies have covered also the success aspects and lessons learned in implementation and development of health IT systems. In many cases evaluation has been led by research interests to develop methodologies or to study the health care processes. From 1980's onwards also management issues, user acceptance and adoption of health IT systems in health care organizations have been studied. Kaplan and Shaw made a review of how aspects related to people, organizational and social issues have been considered in health IT evaluations [9]. They emphasized the need to pay more attention to these issues during

---

<sup>3</sup> See also: J. Brender, Theoretical basis of health IT evaluation, in: E. Ammenwerth, M. Rigby (eds.), *Evidence-Based Health Informatics*, Stud Health Technol Inform 222, IOS Press, Amsterdam, 2016.

<sup>4</sup> See also: D. Luzzi et al., *Economic evaluation of health IT*, in: *ibid.*

<sup>5</sup> See also: L. Lee et al., *Understanding stakeholder interests and perspectives in evaluations of health IT*, in: *ibid.*

health IT system design, implementation and use and emphasize the need to integrate multi-method evaluation to the whole life cycle of the health IT system.<sup>6</sup>

Van der Loo analysed evaluation studies published between 1967 and 1995 with regard to type of the system, study design, data collection, economic evaluation and type of effect measure [10]. He found performance of users to be a criterion in many studies, as well as time savings and costs of patient care. By contrast, user satisfaction was an evaluation criterion only in 11% of studies. These results reflect well the situation in 1990's, usability and effects and impacts of health IT on health care outcomes and quality were not highly valued issues that time.

Ammenwerth and de Keizer [11] found that explanatory research and quantitative methods have dominated evaluation studies during the last 20 years. However, the studies on outcome quality and costs of patient care, patient satisfaction and patient behaviour have received more attention recently as well as studies on quality of processes. At the same time, the number of laboratory studies and technical evaluation studies has declined while the number of studies focusing on the influence of health IT on quality of care processes or outcome of patient care has increased. It was typical for early studies, e.g. evaluation of decision support systems and expert systems, that the focus was on technical issues, on hardware and software quality and on system performance issues. Rigby [12] also noted that the focus of evaluation of a health IT system changes during its life cycle: During the implementation phase evaluation addresses often technical aspects, but with a completed system focuses on impacts on patient care. Ammenwerth and de Keizer found also that the number of inter-organisational studies has increased reflecting the trend towards cooperative and shared care. In many recent studies, user satisfaction and efficiency of patient care are the most frequently addressed evaluation criteria, and there is a slowly growing trend towards evaluation studies covering more than one evaluation criterion [11]. Also the health IT developments discussed in [12,13] showed the importance of evaluation, the challenges and the problems in integrated methodologies and in endorsing the importance of robust studies.

These reviews show that explanative research and quantitative methods have dominated health IT evaluation research for a long time, but studies focusing on process quality or outcome quality of patient care have increased lately. Also human, social and organizational aspects have been included in evaluation studies to some extent, and promisingly research in this aspect is growing [9].<sup>7</sup> Qualitative methods have been used in studies focused on user acceptance, usability and usefulness or on organizational and social impacts of health IT. Many evaluations have been focused on one user's individual context: how a health professional uses the system in the specific context, how he/she accepts the technology and how technology fits into his/her work processes, leaving the more wide usage contexts untouched.

The reviews also reveal quality problems in evaluation studies: there are problems related e.g. to weak study planning without a systematic, scientific methodology, false implicit assumptions made in the study, experimental errors in the research setting, under- or over-interpretation of the results or false conclusions, inclusion of non-

---

<sup>6</sup> See also: P.J. Scott et al., Mixed methods: a paradigm for holistic evaluation of health information system, in: E. Ammenwerth, M. Rigby (eds.), *Evidence-Based Health Informatics*, Stud Health Technol Inform 222, IOS Press, Amsterdam, 2016.

<sup>7</sup> See also: B. Kaplan, *Evaluation of people and organizational Issues – Sociotechnical ethnographic evaluation*, in: *ibid*.

neutral evaluators, intra- or inter-organizational variability in the evaluation object or novelty of technology. In some cases there seem also to be problems in selecting such evaluation methods that are capable to measure those variables and aspects that describe the phenomenon under study. These quality problems show that there is real need to improve the quality of evaluation studies.

Since the early 2000's researchers in the health informatics field have pointed out the need to include also user perspective and usability, in the evaluation studies [14, 15]. This need is motivated by recognition of the importance of good usability as the health IT systems are planned for use in health care clinical practice.<sup>8</sup> We analyse briefly also the published usability evaluation studies, on their methodological approach and quality problems.

A review in [16] showed that in general, empirical usability studies are heavily affected by the traditional approaches to evaluation of human-computer interaction. Usability studies have mainly applied traditional usability evaluation methods, particularly usability testing and inspections. A considerable number of usability evaluation studies have concentrated on the later phases of health IT development, and evaluation has been focused on systems that are already in use. Recent usability studies are characterized by a narrow focus on user and usability issues, emphasis on summative evaluation rather than on design or development, isolated system development, and emphasis on information systems and data management [16]. Also in many usability studies there is a lack of understanding of the contextual aspects of usability, and the characteristics of clinical work contexts, though the widely known definitions for usability [17,18] emphasize the need to understand usability as a contextual property.

There are quality problems also in the usability evaluation studies, e.g. the studies often focus on a single end-user group perspective, user interface components, or use of the system in a specific context, but do not provide a comprehensive picture of usability of a large-scale healthcare information system. Further, the evaluation studies rarely discuss the relationship between single-system development and the existing technology setting in healthcare, or the characteristics of various use contexts in which the evaluated system is used. Typically the studies discuss summative usability results on working systems; this leads to a focus on the problems with adopting current systems in a given healthcare environment and diverts concentration from the design or development of new, better systems. In general, the quality of usability evaluation studies could be enhanced by applying appropriate usability evaluation methods other than traditional ones, e.g. contextual inquiry [19], which takes into consideration the contextual aspects of healthcare environment. The planning of evaluation studies should aim at providing evaluation results but also understanding of the underlying problems and support for the development work.

Usability evaluation studies have in most cases been done separately from other evaluation studies, e.g. from effectiveness, effects and impacts evaluation studies. However, it would be beneficial to integrate usability and other evaluation criteria in a study. Usability is an important quality aspect of health IT and usability is heavily dependent on good health IT system design, implementation and adoption [20]. Some of the recent usability studies have already emphasized a more holistic view for

---

<sup>8</sup> See also: R. Marcilly et al., From usability engineering to evidence-based usability in health IT, in: E. Ammenwerth, M. Rigby (eds.), *Evidence-Based Health Informatics*, Stud Health Technol Inform 222, IOS Press, Amsterdam, 2016.

usability evaluation and called for multi-method approaches and focus more on health professionals' clinical collaboration. Among others, Kushniruk [21, 22] has expressed a concern that when the technology applications become more complex, evaluation methodologies will need to be continually refined in order to keep pace. Multi-method approaches are called for as opposed to using one single method [23]. Kaipio emphasizes [16] that the user-centred approach will play an essential role in future health IT design and development, because knowledge and understanding of the needs of various user groups, as well as clinical use contexts, are important and should be part of all health IT design and development phases.

These findings are supported in [24] by emphasizing the need for predictive evaluation methods to accurately identify usability issues that arise from the interaction, sharing and communication requirements of clinical work. Furthermore, health IT systems have many user groups and the users use the systems for various purposes, such as clinical, nursing, medical, administrative, managerial, statistical and economical purposes and today, also patients and citizens use health IT systems for their personal health and wellness purposes. Therefore, it is important to consider the variety of uses and user contexts, and also the variety of evaluation criteria, since the nature of clinical work, as well as the physical and organizational environments in the workplace, may differ significantly between healthcare units.

This state of affairs of evaluation studies and quality challenges emphasizes the need for systematic approaches and guidelines to design and to carry out different kinds of evaluation studies to provide evidence about the impacts and actual efficiency, quality, usability and safety of health IT [25].

### **3. Guidelines and frameworks for health IT evaluations**

#### *3.1. Overview on available guidelines and frameworks*

Many frameworks and guidelines developed for evaluation exist, aiming at supporting and improving studies so that health IT evaluation is conducted to the highest methodological and scientific standards. These frameworks and guidelines differ in terms of generality, specificity and timing related to system development phases and theoretical underpinning. In this section, we will introduce some selected frameworks and guidelines for health IT evaluation.

Kaplan has suggested an evaluation framework of 4C - Communication, Control, Care and Context [26]. This 4C model calls for multi-method longitudinal design of formative and summative evaluations.<sup>9</sup> Shaw has introduced the CHEATS-framework [27] which identifies six aspects to be important in evaluation: Clinical, Human and Organizational, Educational, Administrative, Technical and Social.

A Model for ASsessment of Telemedicine applications (MAST) [28] lists aspects of evaluation within seven domains of outcomes: health problem and characteristics of the application; safety; clinical effectiveness; patient perspectives; economic aspects; organizational aspects; and socio-cultural, ethical and legal aspects. MAST is planned to be a toolkit, a checklist of issues that need to be considered in evaluation. MAST is

---

<sup>9</sup> See also: B. Kaplan, Evaluation of people and organizational Issues – Sociotechnical ethnographic evaluation, in: E. Ammenwerth, M. Rigby (eds.), *Evidence-Based Health Informatics*, Stud Health Technol Inform 222, IOS Press, Amsterdam, 2016.

based on a health technology assessment (HTA) approach and focuses on telemedicine systems, not on health IT generally. MAST does not consider the execution and management of an evaluation project.<sup>10</sup>

Cusack et al. developed the AHRQ toolkit [29] to provide step-by-step guidance for developing evaluation plans for health IT projects. AHRQ assists evaluators in defining the goals for evaluation, in identifying what is important to stakeholders, what needs to be measured to satisfy stakeholders, what is realistic and feasible to measure, and how to measure these items. The AHRQ toolkit is very useful from the methodological point of view. It can be applied within other more generic guidelines. The toolkit does not, however, give guidance on the evaluation project itself, how to manage it, how to carry out the project, or how to complete and report the study.

A life-cycle framework for evaluation by Clarke et al. [30] is focused on how to evaluate health IT interventions while the system is being designed, developed and deployed. The model is formative and relates evaluation to the phases of the system development. The life-cycle evaluation framework is a valuable tool to monitor the development process and the deployment of a new system.

The HOT-fit evaluation framework [31] considers Human, Organization and Technology factors and recognizes interrelated dimensions of health IT success and determines both benefits and satisfaction. 'Fit' in the framework concerns the ability of health IT system, stakeholders and clinical practices to align with each other.

ISO standards of human-centred design include guidelines for usability evaluation studies [18,32] and state that the following information is needed: Description of the intended goals, a description of the components of the context of use including users, tasks, equipment and environments, target or actual values of effectiveness, efficiency, and satisfaction for the intended contexts. Evaluation should involve the following tasks: Allocating of resources, planning of the evaluation, carrying out sufficient testing, analysing the results and prioritizing issues and proposing solutions, and communicating the solutions appropriately [32].

Many analyses (e.g. [33-34]) have emphasized that we need to understand the changes that health IT systems bring into a complex health care system. The socio-technical assessment tool, STAT-HI [35], focuses on the socio-technical aspects of systems and their implementation in the health care organizational environment.

A recent contribution for health IT evaluation is made by WHO in the context of health IT systems for developing countries [36] (on this topic, see also *Chapter 26: Evaluation of Health IT in low-income countries*). This comprises of nine high-level principles, the Bellagio principles, which cover e.g. evidence-based culture, high quality evaluative data collection and stakeholder engagement (see also *Chapter 1: The need for evidence in health informatics*). The principles emphasize the overall responsibility to ensure that health IT solutions and policies are subjected to, and informed by, rigorous evaluations and that evaluation findings should be used and contribute to evidence generation, synthesis and documentation, including peer-reviewed articles.

---

<sup>10</sup> For details on MAST, see: P. Doupi, *Evolving Health IT Systems Evaluation: The Convergence of Health Informatics and HTA*, in: E. Ammenwerth, M. Rigby (eds.), *Evidence-Based Health Informatics*, Stud Health Technol Inform 222, IOS Press, Amsterdam, 2016.

### 3.2. GEP-HI guideline to design and execute an evaluation study project

The discussed frameworks and guidelines for evaluation aim at improving the quality of health IT evaluation studies. They provide guidance and support on how to plan and perform evaluation studies, but their application scopes and contexts vary as well as theoretical foundations and methodological approaches. Thus a comprehensive evaluation guideline was developed to support especially planning and execution of an evaluation study to overcome the quality problems related to weak study planning and methodological choices. The guideline was named GEP-HI, Good Evaluation Practice Guideline for Health Informatics [7].

The starting point for the GEP-HI guideline development was the existing knowledge, experience and literature on evaluation studies, methodologies, guidelines development, codes of ethics and good implementation practices. In particular the following review materials and textbooks provided the foundation for preparation of the guideline [1, 9, 37-39].

The GEP-HI guideline has been developed through an informal consensus-seeking process, without balloting, in the community of health IT evaluation experts, and it has been regularly in open discussion through the HISEVAL website (<http://iig.umat.at/efmi>) and many conference workshops. The primary authors of GEP-HI were all participants of the ESF HIS-EVAL Workshop [4, 5] and active in the EFMI (European Federation for Medical Informatics) and IMIA (International Association of Medical Informatics) working groups dealing with evaluation of health IT systems (for the details on these working groups, see the book appendix).

The objective of the GEP-HI is to give advice on how to design and carry out evaluation studies in various health IT contexts [7]. The guideline lists issues to consider, and gives recommendations on how to design evaluation studies, how to make methodological choices, how to conduct studies and how to define evaluation criteria at specific phases of the health IT life cycle.

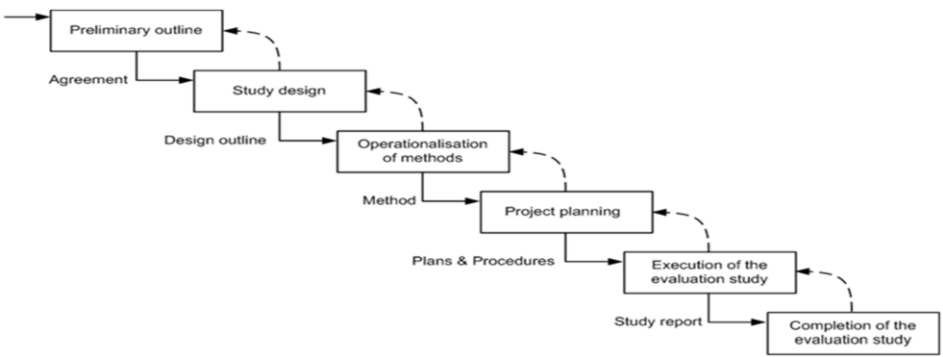
When applied, the GEP-HI guideline has potential to raise the quality of evaluation studies through careful planning, and thus contribute to the accumulation of the scientific evidence base. GEP-HI is complemented by the STARE-HI – Statement on Reporting of Evaluation Studies in Health Informatics [38], which provides guidance on how to report an evaluation study (for details on STARE-HI, see *Chapter 24: Publishing health IT evaluation studies*).

The GEP-HI guideline [7] is divided into parts corresponding to the phases of an evaluation study (Figure 1). The theoretical background for the study phases is compliant with the information system development models. Implementation is an iterative spiral; the topics are in general repeated in depth or breadth to achieve progress during all phases, and feedback loops urge to revisit earlier phases when new aspects, additional information, or changes in context appear.

The phases of GEP-HI guideline are:

- *Preliminary outline* presenting the purpose of the study and the first ideas on why, for whom, and how the evaluation should take place,
- *Study design* clarifying the design issues for the evaluation study.
- *Operationalization of methods* making the methodological approach and methods concrete and compliant with the system type, the organization and the information need.
- *Project planning* developing plans and procedures for the evaluation project.

- *Execution of the evaluation study* accomplishing the designed evaluation study.
- *Completion of the evaluation study* reporting, accounting, archiving of evaluation study results, finalization of outstanding issues and formal closure of the evaluation study.



**Figure 1:** Phases of a health IT evaluation according to the GEP-HI guideline [7].

To progress from one phase to the next phase a formal acceptance is required from the relevant stakeholders of the planned evaluation study. For each phase a list of items is presented (Table 1) and these should be carefully considered during the study. All phases together contain some 60 detailed items, which are presented in relation to the evaluation study phases. When designers and executers of evaluation studies address these items, the plan, structure, objectives and results of the studies will become more robust and consequently the studies contribute an important step towards evidence-based health informatics.

**Table 1.** Items to be considered at each phase of a health IT evaluation according to GEP-HI.

Phase no.	Phase	Items of the phase
1	Preliminary outline	<ul style="list-style-type: none"><li>- Purpose of the study</li><li>- Primary audience</li><li>- Identification of the study funding party(ies)</li><li>- First identification of stakeholders</li><li>- Identification of required expertise</li><li>- The organizational and user context of the evaluation study</li><li>- Object of evaluation, type of health IT</li><li>- First exploration of evaluation methods to be used</li><li>- Ethical and legal issues</li><li>- Budget</li><li>- Preliminary permissions for publication</li><li>- Result of preliminary outline</li><li>- Formal acceptance to proceed to the next phase</li></ul>



2	<b>Study design</b>	<ul style="list-style-type: none"> <li>- Detailed rationale and objectives for the study</li> <li>- Key evaluation issues, questions, indicators</li> <li>- Stakeholder analysis/Social Network analysis</li> <li>- Study methods</li> <li>- Organizational context, the study setting</li> <li>- Technical setting, the type of health IT</li> <li>- Participants from the organization</li> <li>- Project timeline</li> <li>- Material and practical resources</li> <li>- Establishment of the study team</li> <li>- Risk analysis and quality management</li> <li>- Budget</li> <li>- Ethical and legal issues</li> <li>- Strategy for reporting and disseminating the results</li> <li>- Result of study design</li> <li>- Formal acceptance to proceed to the next phase</li> </ul>
3	<b>Operationalization of methods</b>	<ul style="list-style-type: none"> <li>- Study type</li> <li>- Approach</li> <li>- Assumptions and feasibility assessment</li> <li>- Frame of reference</li> <li>- Timing</li> <li>- Justification of the methodological approach</li> <li>- Expertise</li> <li>- Outcome measures</li> <li>- Avoiding Bias</li> <li>- Quality control on data (measures)</li> <li>- Participants</li> <li>- Ethical and legal issues</li> <li>- Result of operationalization of methods</li> <li>- Approval of operationalization of methods</li> </ul>
4	<b>Project planning</b>	<ul style="list-style-type: none"> <li>- Project management</li> <li>- Study flow</li> <li>- Evaluation activity mapping</li> <li>- Quality management</li> <li>- Risk management</li> <li>- Recruitment of necessary staff</li> <li>- Inform all relevant stakeholders</li> <li>- Result of project planning</li> <li>- Approval of project planning</li> </ul>
5	<b>Execution of the evaluation study</b>	<ul style="list-style-type: none"> <li>- Undertake the study, collect data and interpret observations</li> <li>- Quality control of findings and observation of changes</li> <li>- Continuous project management, quality and risk management</li> <li>- Regular reports</li> <li>- Final result of execution of the evaluation study</li> </ul>
6	<b>Completion of the evaluation study</b>	<ul style="list-style-type: none"> <li>- Accounting</li> <li>- Archiving</li> <li>- Reports and publications</li> </ul>

The strength of the GEP-HI guideline is in forcing the user to go through a checklist of relevant issues that might otherwise only act informally as tacit knowledge, or even be overlooked. This systematic approach will increase the likelihood of an evaluation outcome with the desired level of accuracy and precision and hence an increased effectiveness, and additionally encourage the adoption of a scientifically

valid approach in an evaluation study. The performed evaluation study should be reported following the STARE-HI reporting statement [38].<sup>11</sup>

#### **4. Discussion and conclusions**

The scope of evaluation studies is wide. There are small snapshot-type studies just to get an insight on the ongoing development, or longitudinal wide studies following the health IT system for a long time with various users and use contexts, or impact studies to assess the changes that have been implemented by the health IT system and possibly also measure the effects of the system on efficiency and outcomes of the health care organization. All these require different planning, management and methods, different resources and expertise.

There are quality problems in reported evaluation studies, e.g. weak planning, missing systematic, scientific methodology for evaluation, false assumptions and conclusions, experimental errors and weak attention to intra- or inter-organizational variability. Quality problems in usability evaluations are related e.g. to the narrow focus of the study, or evaluation of the use in a very specific context, or leaving out the relationship between single-system development and the existing technology setting in healthcare or the characteristics of various use contexts.

In health care clinical context, the environment consists of many health IT applications, of which several are used simultaneously. Evaluation of these health IT environments should address the relevant evaluation criteria from a broad viewpoint. Health IT systems need to be seen as integrated parts of a wider technology environment and the objectives of evaluation should be framed with respect to the clinical situation and use contexts. A challenge is to cover the wide variety of users of the health IT systems and the numerous purposes these systems serve, and the diversity of clinical surroundings in healthcare organizations where the systems are implemented and used.

The GEP-HI guideline has been developed to overcome the identified quality problems in evaluation study design and execution. GEP-HI can be applied to different kinds of health IT evaluation studies, irrespective of whether the object of study is an IT application or a method like nursing classification or data security practice. In small evaluation studies not all phases of the GEP-HI guideline may be needed. The guideline is applicable at various phases of a health IT project, starting from design and development, over application or system implementation and installation, and ending with the study of effects and impacts in routine use.

The GEP-HI guideline and other discussed health IT evaluation guidelines and frameworks support detailed planning of the evaluation study and thus help to consider carefully the study objectives and operationalisation of methods for the specific evaluation study in planning; most of them are applicable to different types of studies, such as feasibility, effectiveness, efficiency and impact evaluation, and to studies with various scopes. GEP-HI supports application of multi-method approach in evaluation and integration of usability evaluation with the other evaluation criteria e.g. interoperability, security, effects and impacts evaluation.

---

<sup>11</sup> See also: P. Doupi, *Evolving Health IT Systems Evaluation: The Convergence of Health Informatics and HTA*, in: E. Ammenwerth, M. Rigby (eds.), *Evidence-Based Health Informatics*, Stud Health Technol Inform 222, IOS Press, Amsterdam, 2016.

Rigby et al. [40] listed the ten core principles that are essential for the effectiveness in all evaluations: Preliminary planning, stakeholder analysis, health issue and eHealth application, safety, clinical effectiveness, user experience, economic aspects, organisational aspects, ethical and legal issues and reporting of the studies. These are well covered by the GEP-HI guideline. For the future, it is important to define the quality aspects of interest in an evaluation study and to plan the study carefully and apply approaches that enable integration of various evaluation criteria in the specific evaluation study. This is needed to get a complete picture of the health IT system under study, from the users', developers' and from health care organisational perspectives. Systematic guidelines help to pay attention to all relevant issues and to plan carefully the evaluation study and execute the study following scientific principles of research.

### **Recommended further readings**

1. E. Ammenwerth, J. Brender, P. Nykänen, H.U. Prokosch, M. Rigby, J. Talmon, Visions and strategies to improve evaluation of health information system. Reflections and lessons based on the HIS-EVAL workshop in Innsbruck, *Int J Med Inform* **73** (2004), 479-91.
2. E. Ammenwerth and N. de Keizer, An inventory of evaluation studies of information technology in health care. Trends in evaluation research 1982 – 2002, *Meth Inf Med* **1** (2005), 44-56.
3. C.F. Friedman, J. Wyatt, *Evaluation Methods in Medical Informatics*, 2nd edition, Springer, New York, 2006.
4. M. Rigby, E. Ammenwerth, M-C. Beuscart-Zephir, J. Brender, H. Hyppönen, S. Melia, P. Nykänen, J. Talmon, N. de Keizer, Evidence based health informatics: 10 years of efforts to promote the principle, in: *IMIA Yearbook of Medical Informatics 2013*, Schattauer, Stuttgart, 2013. pp. 1-13.

### **Food for thought**

1. What are the challenges and possibilities to integrate various evaluation frameworks or guidelines in one evaluation study? Think e.g. integration of GEP-HI and MAST or integration of GEP-HI and HOT-fit?
2. How to integrate usability evaluation with standard evaluation framework, e.g. within the GEP-HI guideline?
3. Analyse the differences between a small-scale and a wide-scale evaluation study planning and execution. Differences may be related to e.g. objectives of the study, scope of the study, evaluation criteria, methods applied, reporting of the study. As examples of various study scopes you may consider e.g. evaluation of the user acceptance of one health IT system in one hospital department (small-scale) and evaluation of the impacts of an electronic health record system (EHR) at many levels of the health care system (primary care, specialized care, tertiary care).

## References

- [1] J. Brender, *Handbook of evaluation methods for health informatics*, Elsevier Academic Press, USA, 2006.
- [2] E. Ammenwerth, N. Shaw, Bad health informatics can kill – Is evaluation the answer, Editorial. *Methods Inf Med* **44** (2005), 1-3.
- [3] A.G. Ekeland, A. Bowes, S. Flottorp, Effectiveness of telemedicine: A systematic review of reviews, *Int J Med Inform* **79** (2010), 736-771.
- [4] E. Ammenwerth, J. Brender, P. Nykänen, H.U. Prokosch, M. Rigby, J. Talmon, Visions and strategies to improve evaluation of health information systems. Reflections and lessons based on the HIS-EVAL workshop in Innsbruck, *Int J Med Inform* **73** (2004), 479-91.
- [5] *Declaration of Innsbruck – Results from the European Science Foundation sponsored Workshop on Systematic Evaluation of Health Information Systems (HIS-EVAL)*, April 4-6th, 2003. Published in: Ammenwerth E, Brender J, Nykanen P, Prokosch HU, Rigby M, Talmon J, et al., Visions and strategies to improve evaluation of health information systems; Reflections and lessons based on the HIS-EVAL workshop in Innsbruck, *Int J Med Inf* **73**(6) (2004), 479-91.
- [6] M. Rigby, *Evaluation: 16 Powerful Reasons Why Not to Do It - And 6 Over-Riding Imperatives*, in: V. Patel, R. Rogers, R. Haux (eds.), *Medinfo 2001: Proceedings of the 10th. World Congress on Medical Informatics*, IOS Press, Amsterdam, 2001, pp. 1198-1202.
- [7] P. Nykänen, J. Brender, J. Talmon, N. de Keizer, M. Rigby, M-C. Beuscart-Zephir, E. Ammenwerth, Guideline for good evaluation practice in health informatics, *Int J Med Inform* **80** (2011), 815-827.
- [8] E.M.S.J. van Gennip, J.T. Talmon (eds.), *Assessment and evaluation of information technologies*, IOS Press, Amsterdam, 1995.
- [9] B. Kaplan, N. Shaw, Future directions in evaluations research: People, organisational and social issues, *Methods Inf Med* **43** (2004), 215-231.
- [10] R. van der Loo, *Overview of Published Assessment and Evaluation Studies*, in: E.M.S.J. van Gennip, J.T. Talmon (eds.), *Assessment and evaluation of information technologies*, IOS Press, Amsterdam, 1995. pp. 261-282.
- [11] E. Ammenwerth, N. de Keizer, An inventory of evaluation studies of information technology in health care, *Trends in evaluation research 1982 – 2002*, *Meth Inf Med* **1** (2005), 44-56.
- [12] M. Rigby, *Evaluation – the Cinderella science of ICT in health*, in: C. Kulikowski, R. Haux (eds.), *IMIA Yearbook of Medical Informatics 2006*, Schattauer, Stuttgart, 2006. pp. 114-120.
- [13] J. Talmon, *Evaluation and implementation: A call for action*, in: C. Kulikowski, R. Haux (eds.), *IMIA Yearbook of Medical Informatics 2006*, Schattauer, Stuttgart, 2006. pp. 11-15.
- [14] J. Zhang, Human-Centred Computing in Health Information Systems, Part 1: Analysis and Design, *Journal of Biomedical Informatics* **38** (2005), 1-3.
- [15] B. Chaudhry, J. Wang, S.Wu, M. Maglione, W. Mojica, E. Roth, S.C. Morton, P.G. Shekelle, Systematic Review: Impact of Health Information Technology on Quality, Efficiency, and Costs of Medical Care, *Annals of Internal Medicine* **144** (2006), 742-752.
- [16] J. Kaipio, *Usability in healthcare: Overcoming the mismatch between information systems and clinical work*, Aalto University publication series, Doctoral dissertations 105, Unigrafia Oy, Helsinki, 2011, <http://lib.tkk.fi/Diss/2011/isbn9789526043340/isbn9789526043340.pdf>, last access 11 February 2016.
- [17] J. Nielsen, *Usability Engineering*, Academic Press, San Diego, USA, 1993.
- [18] ISO 9241-11, *Ergonomic Requirements for Office Work with Visual Display Terminals, part 11: Guidance on Usability*, International Organization for Standardization, Geneva, 1998.
- [19] H. Beyer, K. Holtzblatt, *Contextual Design: Defining Customer-Centered Systems*, Academic Press, San Diego, USA, 1998.
- [20] M-C. Beuscart-Zephir, P. Elkin, S. Pelayo, R. Beuscart, *The human factors engineering approach to biomedical informatics projects: State of the art, results, benefits and challenges*, in: A Geissbuhler, R Haux, C Kulikowski (eds.), *IMIA Yearbook of Medical Informatics 2007*, Schattauer, Stuttgart, 2007. pp. 109-127.
- [21] A. W. Kushniruk, Evaluation in the Design of Health Information Systems: Application of Approaches Emerging from Usability Engineering, *Computers in Biology and Medicine* **32** (2001), 141-149.
- [22] A.W. Kushniruk, V.L. Patel, Cognitive and usability engineering methods for the evaluation of clinical information systems, *J Biomed Inform* **37**(2004), 56-76.
- [23] M.W.M. Jaspers, A comparison of usability methods for testing interactive health technologies: Methodological aspects and empirical evidence, *Int J Med Inform* **78** (2009), 340-353.
- [24] P.J. Edwards, K.P. Moloney, J.A. Jacko, F. Sainfort, Evaluating usability of a commercial electronic health record: a case study, *Int J Hum Comput Stud* **66** (2008), 718-728.
- [25] E. Ammenwerth, S. Gräber, G. Herrmann, T. Bürkle, J. König, Evaluation of Health Information Systems – Problems and Challenges, *Int J Med Inform* **71** (2003), 125–135.

- [26] B. Kaplan, Addressing organizational issues into the evaluation of medical systems, *J Am Med Inform Assoc* **4** (1997), 94-101.
- [27] N.T. Shaw, CHEATS: A generic evaluation framework for information communication technologies, *Comp Biol Med* **32** (2002), 209-220.
- [28] K. Kidholm, A. Bowes, S. Dyrehauge, A. Granström Ekeland, S.A. Flottorp, L. Kvistgaard et al., *The MAST Manual - MAST-Model for assessment of telemedicine*, <http://www.epractice.eu/en/library/315610>, last access 11 February 2016.
- [29] C.M. Cusack, C. Byrne, J.M. Hook, J. McGowan, E.G. Poon, A. Zafar, *Health information technology evaluation toolkit: 2009 Update*, AHRQ Publication No. 09-0083-EF, Agency for Healthcare Research and Quality, <http://www.ahrq.gov>, last access 11 February 2016.
- [30] K. Clarke, R. O'Moore, R. Smeets, J. Talmon, J. Brender, P. McNair et al., A methodology for evaluation of knowledge-based systems in medicine, *Artificial Intelligence in Medicine* **6** (1994), 107-121.
- [31] M.M. Yusof, J. Kuljis, A. Papazafeiropoulou, K. Stergioulas, An evaluation framework for health information systems: Human, organization and technology-fit factors (HOT-fit), *Int J Med Inform* **77** (2008), 386-398.
- [32] ISO 9241-210 (2010) *International Standard: Ergonomics of Human-system Interaction – Part 210: Human-centred Design for Interactive Systems*, First edition 2010-03-15, Reference number ISO 9241-210:2010(E).
- [33] M.M. Yusof, A. Papazafeiropoulou, R.J. Paul, K. Stergioulas, Investigating evaluation frameworks for health information systems, *Int J Med Inform* **77** (2008), 377-385.
- [34] M. Berg, Medical work and the computer-based patient record: a sociological perspective, *Meth Inf Med* **7** (1998), 294-301.
- [35] P.J. Scott, J. Briggs, STAT-HI: A socio-technical assessment tool for health informatics implementations, *Open Med Inform J* **4** (2010), 214-20.
- [36] WHO, *Call to Action on Global eHealth Evaluation - Consensus Statement of the WHO Global eHealth Evaluation Meeting*, Bellagio, [http://www.ghdonline.org/uploads/The\\_Bellagio\\_eHealth\\_Evaluation\\_Call\\_to\\_Action-Release.docx](http://www.ghdonline.org/uploads/The_Bellagio_eHealth_Evaluation_Call_to_Action-Release.docx), last access 11 February 2016.
- [37] C.F. Friedman, J. Wyatt, *Evaluation Methods in Medical Informatics*, 2nd edition, Springer, New York, 2006.
- [38] J. Talmon, E. Ammenwerth, J. Brender, N. de Keizer, P. Nykänen, M. Rigby, STARE-HI – Statement on Reporting of Evaluation Studies in Health Informatics, *Int J Med Inform* **78** (2009), 1- 9.
- [39] P. Nykänen, J. Enning, J. Talmon, D. Hoyer, F. Sanz, C. Thayer, R. Roine, M. Vissers and F. Eurlings, Inventory of validation approaches in selected health telematics projects, *Int J Med Inform* **56** (1999), 87-96.
- [40] M. Rigby, E. Ammenwerth, M-C. Beuscart-Zephir, J. Brender, H. Hyppönen, S. Melia, P. Nykänen, J. Talmon, N. de Keizer, *Evidence based health informatics: 10 years of efforts to promote the principle*, in: IMIA Yearbook of Medical Informatics 2013, Schattauer, Stuttgart, 2013. pp. 1-13.